Implementing RSVP as an Image Browser

Andrew Barrett
Department of Computer Science
University of Canterbury
Christchurch, New Zealand

Supervisor: Andy Cockburn

Honours Report - 2002

Abstract

Rapid Serial Visual Presentation (RSVP) is an effective tool for information navigation. We applied RSVP to the increasingly important task of image browsing. In total nine RSVP interfaces were developed, implementing three RSVP modes and three control sets. An evaluation was carried out to investigate the effect of RSVP mode and control set on image retrieval time. Subjective measures were also taken using the NASA Task Load Index (TLX) worksheets. In the evaluation, a traditional thumbnailed image browser was used, so as to give a base for comparison. Choice of RSVP mode was found to have the greatest effect on image retrieval speed and user preference, with interfaces implementing floating RSVP performing significantly better than those implementing other RSVP modes. The floating RSVP interfaces were faster than the thumbnailed approach but not significantly so. Subjective satisfaction favoured the floating RSVP interfaces over the thumbnailed interface.

Contents

Chapter 1

Introduction

1.1 Overview

The first digital image was scanned onto a computer in 1957, when Russel A. Kirsch used a drum scanner to convert an image of his baby son into a binary format, which could then be stored on a computer for further processing (Mikkel 1992). The past 50 years have seen a massive increase in both computer processing power and storage. These factors, combined with an increase in the quality and affordability of digital cameras and scanners, has seen the amount of graphical information stored on computers skyrocket.

Computer users are maintaining huge collections of images, both as personal photo collections and as image databases to aid in the compilation of professional computer documents. However, these digital collections are only as useful as the tools available to manipulate them, so it is important to provide a way to quickly search and browse digital image collections.

The amount of digital media stored on computers will continue to increase for the foreseeable future, and thus, the need for effective picture management systems is becoming increasingly more important.

Picture management systems can be split into two main parts: categorisation and search, and image browsing. The categorisation and search function helps to return a subset of images, which the user then browses through. We focus on the image browsing function in our research. An image browser is an application that allows users to select an image from a collection of images. In this paper we investigate several image browsers that utilise an information presentation technique known as Rapid Serial Visual Presentation (RSVP).

RSVP was introduced by Forster in 1970. It was developed as a technique for studying language processing and comprehension (Forster 1970). The technique has since been generalised to the field of information navigation (de Bruijn & Spence 1999).

1.2 Motivation

While many image browsers and several browsing techniques now exist, image browsing is still a young field. Work is needed to ensure that people have good tools with which to make use of their digital information.

Research by de Bruijn & Spence (1999) has suggested that RSVP could be implemented effectively to search and browse electronic information. We were interested in conducting experiments to test the possibility that an effective image browser could be implemented using RSVP. To this effect we implemented several RSVP modes that have previously been proposed and came up with three new techniques for controlling the RSVP process. We judged effectiveness by the time taken to retrieve an image and user satisfaction.

The report contributes a valuable evaluation of several RSVP modes and presents several new ways with which to control RSVP.

1.3 Report Structure

Chapter 2 provides an overview of previous work in the fields of RSVP and image browsing. Chapter 3 discusses the design decisions behind the interfaces. Chapter 4 describes the evaluation conducted to assess the effectiveness of the interfaces. Chapter 5 discusses the results from the evaluation and presents some suggestions for further work. Chapter 6 concludes the report.

Chapter 2

Related Work

This chapter is split into two sections: the previous work done on image browsing, and that done on RSVP.

2.1 Image Browsing

While there are many image browsing applications available for download from the Internet, the majority implement a thumbnailed folder approach to image browsing. This approach displays a grid of thumbnailed pictures that can be browsed through with a scroll bar. ACDSee is a popular image browser that utilises this approach, and this is shown in Figure 2.1.

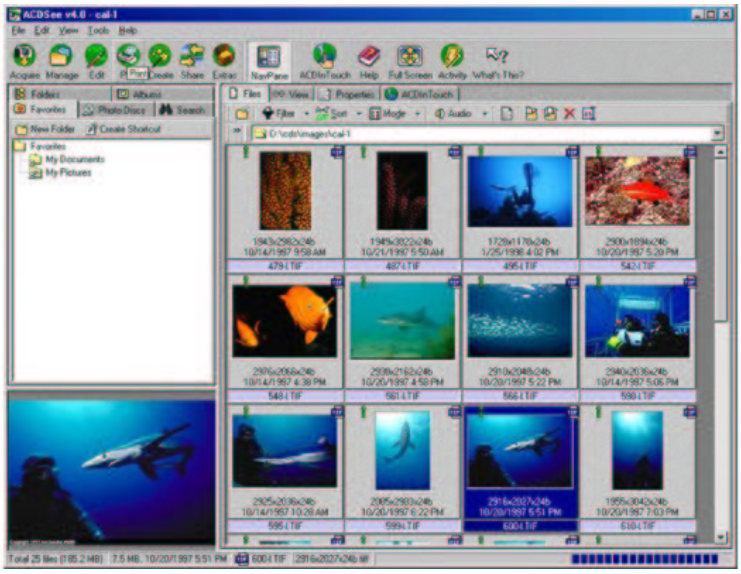


Figure 2.1: ACDSee Thumbnail Image Browser

Most of the research done on improving image browsing has been aimed at improving the thumbnailed approach as opposed to developing new interfaces. A zoomable image browser (ZIB) was developed by Combs & Bederson (1999), which allowed users to change the size of the thumbnails shown and therefore allow more or fewer pictures on screen at any one time . This technique was found not to be faster than the traditional thumbnailed approach. Bederson (2001) also developed an image browser that utilised Zoomable User Interface technology. Directories of photos are represented as part of a treemap, with users able to zoom in on a region of photos until they are presented with a single image. The system they developed was named Photomesa. While is was reported that Photomesa helped users to comfortably

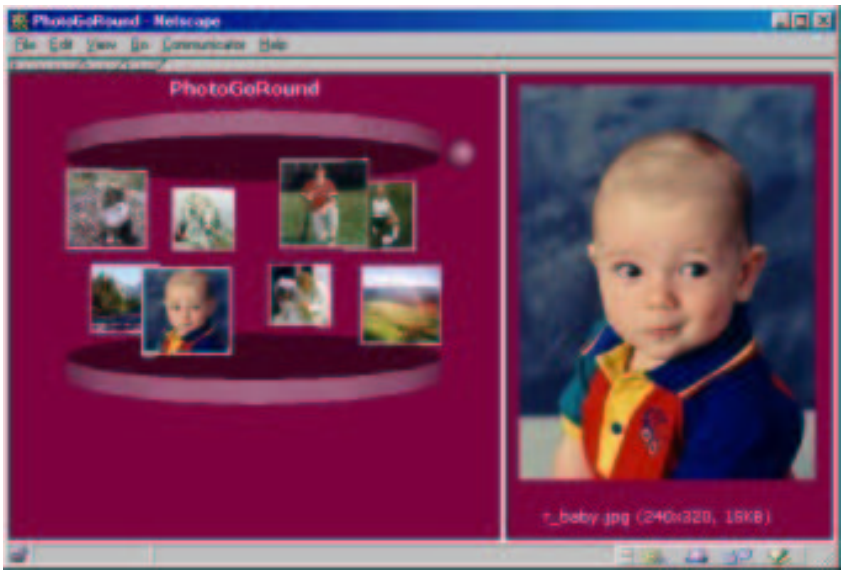


Figure 2.2: Photo GoRound Image Browser

browse large image sets, they provide no formal results.

In the study by Combs & Bederson, ZIB was also tested against two non-traditional browsers developed by the TriVista corporation (TriVista Technologies 2002); Simple LandScape and Photo GoRound. Figure 2.2 shows the Photo GoRound browser. Both of TriVista's browsers performed poorly in terms of time taken to retrieve an image. However, the users did find Photo GoRound the most entertaining interface.

Tatemura (2000) took an alternate approach. In a paper he describes an image browser that groups images by shape similarity and then displays them on a Fisheye viewer. This browser is shown in Figure 2.3. Unfortunately, no formal evaluation of the interface was undertaken. However, it remains an interesting technique.

Research has also been conducted on how to organise the images rather than how to present them to the user. Rodden, Basalaj, Sinclair & Wood (2001) investigated whether the efficiency of an image browser is increased by organising thumbnails by mutual similarity. They tested similarity in terms of visual similarity and caption based similarity when the images were manually grouped together. They compared a task where users had to select a target image, and the image set was either randomly sorted or sorted by visual similarity, so that like images appeared next to each other. The comparison was based largely on subjective user measures. They concluded that organisation by similarity seemed to be a useful tool for designers, although more work with different picture organisations needed to be done.

One approach to image organisation is to categorise images by the time that the images were captured. The Calender Browser developed by Graham, Garcia-Molina, Paepcke & Winograd (2002) uses this approach. Users were able to find images 33% faster using the Calender Browser when compared to a commercially available browser. As the images are grouped by time taken, photos from the same events are grouped together, and users were able to use their knowledge of how events related together to find photos. However, this seems dependant on the photos being personal to the user, and would not work if they were browsing through an unfamiliar image collection.

Another way of organising images is by hierarchical clustering. Krishnamachari (1999) developed a system where images are automatically clustered using a hierarchical clustering algorithm, which allowed users to browse through the images by navigating the resulting tree-like structure. Krishnamachari concluded that the hierarchical tree structure can be used to efficiently navigate through large volumes of images; but again, no evaluation of image retrieval time was undertaken.

Figure 2.3: Dynamic Fisheye Image Browser

2.2 RSVP

Rapid Serial Visual Presentation, or RSVP, was introduced by Forster in 1970. It was developed as a method to study language processing and comprehension (Forster 1970). An analogy for the method is a person flicking through the pages of a book scanning for information.

Research by de Bruijn & Spence (1999) suggested that RSVP could be implemented effectively for searching and browsing electronic information. An RSVP mode is the term given to a specific implementation of the RSVP process. RSVP modes include keyhole, carousel, collage and floating RSVP (Spence 2002). Each of these modes rapidly displays a series of images or text.

RSVP has already been evaluated in several contexts as a tool for information navigation. These include an online bookstore by Wittenburg et al. (1998, 2000a, 2000b) and a video selection tool developed by Tse et al. (1998).

de Bruijn & Spence concluded that RSVP is a valuable technique for searching and browsing information on small screen displays, such as PDAs and mobile phones. Research by Rahman & Muter has shown that RSVP can be used to present text on small displays, and is as efficient as the normal page format. Normal page format in analogous to text in the page of a book. Participants in their study showed no preference between RSVP and the normal page format (Rahman & Muter 1999).

To create an RSVP image browser we needed to decide which RSVP mode to use and how to control the process. This is explained further in Chapter 3.

2.2.1 Controls

Most of the work done with RSVP in the field of information navigation focuses on information previewing, rather than browsing. With previewing the information is flashed before the user in order to provide them with an indication of the content of the information. For these systems, controls beyond a simple start and stop button are unnecessary. When used in the text processing domain, RSVP systems typically display words or sentences at a constant rate, like an automatic slide show. Little work has been done on developing

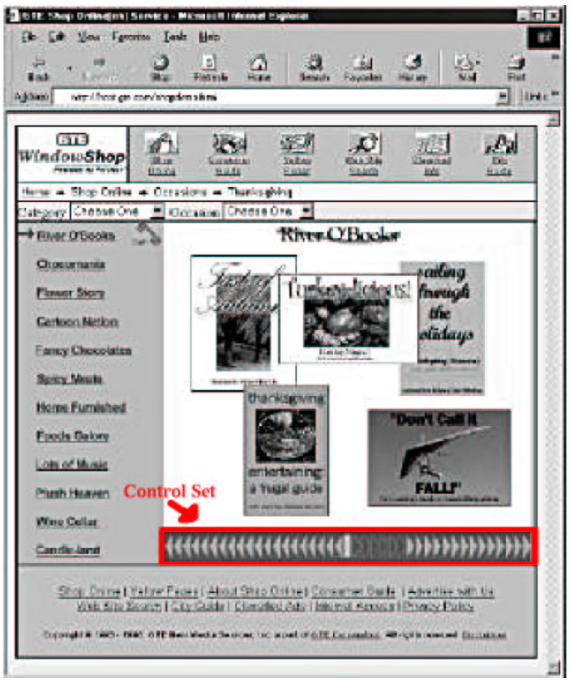


Figure 2.4: Online Bookstore

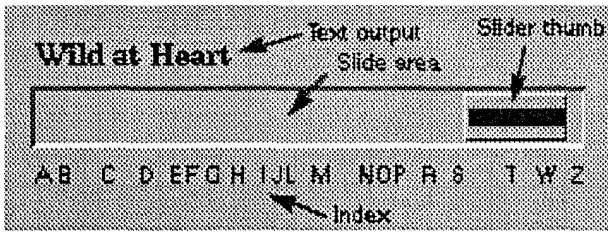


Figure 2.5: The Alphaslider Interface

a full control set for RSVP. We define ‘full control set’ as one that allows users to find and select a piece of information they want from the RSVP display.

Wittenburg et al. (2000b) developed several web-based interfaces that use rapid fire imaging, another term for RSVP. The interface controls were initially those of a typical media player (stop, pause, forward, fast-forward, reverse, fast-reverse). In an evaluation of these controls, users could not stop a sequence of pictures quickly enough to select the required image.

The research indicated that a set of controls for RSVP browsing must support multiple speeds and changes of direction with instant response times. One such set of controls was implemented in an online bookstore browsing system, which is shown in Figure 2.4.

The control set that was used in the online bookstore consisted of a bar of control arrows. Placing the cursor over these arrows started the RSVP process, and moving the cursor off the arrows stopped it. The speed of the RSVP process was controlled by the cursor’s distance from the center of the control arrow set. The direction of the RSVP motion was dictated by what side of the center the cursor lay.

Several list selection tools were developed using RSVP (Ahlberg & Shneiderman 1994). They consisted of a small peice of text output that represented the current selection, and a control set consisting of a slider and is shown in Figure 2.5. It was found that novice users could select an item from a list of 10,000 in 24 seconds and expert users could do so in 13 seconds. No evaluation was done against traditional list

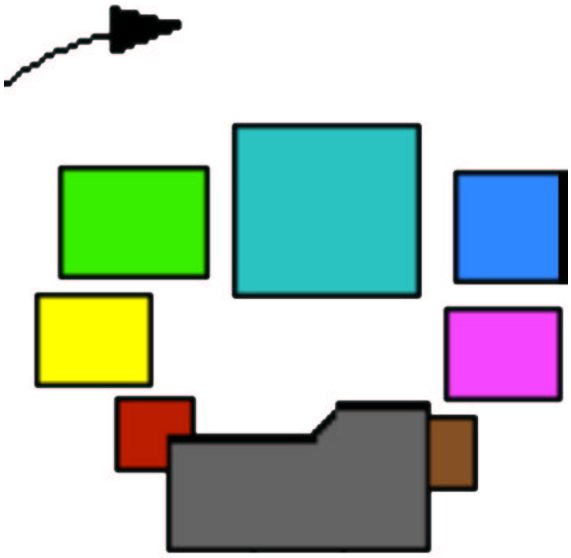


Figure 2.6: Carousel RSVP

selection techniques. Four control sets were used, though they all employed a similar slider interface, with the difference being how the user selected the speed at which the slider could be moved.

2.2.2 RSVP Modes

Of the many RSVP modes, only a few have been formally evaluated, including keyhole, collage, floating and carousel RSVP.

Keyhole RSVP is similar to a slide show. A single image is presented and updated at a regular time interval, always at the same location.

Carousel RSVP displays a series of images on the screen simultaneously. Starting off small, images rotate from the left in a clockwise direction. An image reaches its maximum size at the top of the rotation, and grows smaller as it reaches the end of the rotation at the bottom right. This process is shown in Figure 2.6.

Collage RSVP is a method where images are overlaid on the previous images, similar to pictures being dropped on a table in front of a user (Spence 2002). This is the mode utilised for the online book store, depicted in Figure 2.4.

Floating RSVP is shown in Figure 3.1. When scrolling forward through the image set, the pictures come toward the user with new pictures appearing in the distance and older pictures growing larger until they disappear 'off' the screen. This mode is similar to driving down a highway with billboard advertisements. It is possible to see the billboards in the distance, and if one proves interesting a driver can focus on it until passes by.

de Bruijn and Spence compared the effectiveness of carousel and keyhole RSVP. Subjects looked at a target image for as long as they wished. They were then shown an RSVP display of 20 images and were asked if the target image had been displayed in this set of 20 images. Each participant performed the task seven times with both keyhole and carousel RSVP. When carousel RSVP was used, subjects could accurately report the presence of the target image with presentation times as low as 100 milliseconds. When keyhole RSVP was used, the target image was reported accurately when images were displayed for as little as 40 milliseconds (Spence 2002).

In Wittenburg et al.'s (2000b) online bookstore, both collage RSVP and slide show RSVP were implemented, in addition to a set of conventional HTML pages. The collage mode implementation is shown in Figure 2.4. Users were asked to rate the acceptability of these three techniques from 1 to 5, with 1 being most favoured. The results of this ranking were significant, with the keyhole implementation ranked most highly at 1.37, HTML ranked 2.05 and collage at 2.58.

de Bruijn & Spence (2002) looked at RSVP modes from the point of view of patterns. They studied several RSVP modes including carousel, floating and collage. They found that none of the modes posed any perceptual problem to the user. However, they did suggest that the modes that use moving images, such as carousel and floating, may cause more eye strain than modes that present static images, such as keyhole and collage.

2.2.3 Application to Small Screen Devices

Much of the research undertaken in RSVP has focused on its application to small screen devices. Robert Spence terms it the space time tradeoff, where less information is shown for a shorter period of time (de Bruijn & Spence 1999). RSVP is already used in mobile devices to enhance the reading of text (Wobbrock, Forlizzi, Hudson & Myers 2002). RSVP has also been looked at to aid web browsing in mobile devices (de Bruijn, Spence & Yin 2001). In particular de Bruijn et al. looked at how to answer common web browsing questions such as 'Where am I?', 'Where can I go?' and 'Where have I been?'. They concluded that traditional web browsing methods applied to mobile devices required unacceptable levels of scrolling and that RSVP could successfully be used to overcome this.

Several new mobile phones and PDA's allow the ability to take digital pictures and send them to other mobile devices. Unfortunately traditional thumbnailed image browsers do not perform when the screen real estate is limited (Combs & Bederson 1999). RSVP may provide a means to implementing a useful image browser in a mobile device. Derthick (2002) developed an RSVP interface for an image search on a palmtop to see if it performed better than a traditional approach. The RSVP interface developed did not offer an improvement, and most users disliked using it. However, the implementation used was a fixed rate keyhole with no controls past start and stop. It is our belief that better controls would have allowed users to both perform better and enjoy using the interface more.

Chapter 3

Experimental RSVP Modes & Control Sets

Ten interfaces were constructed for the evaluation. Nine of the interfaces were a combination of three RSVP modes with three control sets. The tenth was a simple thumbnail browser.

The interfaces were written in Tcl/Tk and shared a similar design. The RSVP interfaces comprised of two frames, a RSVP display frame and a control frame. The interfaces were designed to run under any operating system that supports Tcl/Tk and have been tested under Red Hat Linux, Windows XP, Windows 2000 and Windows 98.

3.1 RSVP Interfaces

All nine RSVP interfaces consisted of a display frame and a control frame. The display frame was were the images were displayed and the control frame contained the control set and a counter. The counter was provided to give users feedback into how far through an image set they were. The counter displayed the current image number and the total number of images.

3.1.1 RSVP Modes

The three modes that were chosen all employed quite a different approach. Keyhole RSVP displays a simple static 2D approach with a single image, carousel RSVP displays multiple images in 2D approach with animation and floating RSVP displays multiple images in 3D wit animation. By using the three quite different modes, it was hoped to gain an indication of what is necessary in an RSVP mode to implement a successful image browser.

Keyhole

The keyhole implementation used a standard slideshow approach, with a static image being replaced at a regular interval. The display window was 450x250 pixels, with the image centered in the display. This is shown in Figure 3.3.

We chose the keyhole mode for several reasons. It has been shown to be an effective way to search information spaces (Spence 2002). In addition, Spence hypothesised that keyhole RSVP would minimise eyestrain in users due to the static nature of the image presentation. A possible disadvantage of keyhole interfaces is that as only one image is displayed on screen at a time, there is no chance to spend less time looking at one image in favour of a more interesting image.

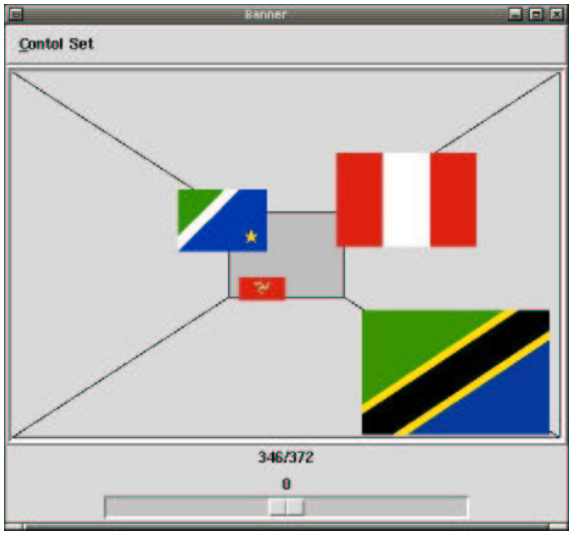


Figure 3.1: Floating RSVP with the Scroll Bar Control Set

Floating

The floating RSVP interface is shown in Figure 3.1. The display window was 450x300 pixels. At any one time four images were present on the screen. Shading and lines were added to create a 3D perspective effect, with smaller images appearing further in the distance than larger images. While the RSVP motion is in the forward direction, new images appear in the distance and are small, they then travel forward toward the user, growing larger as they travel. When images appear to reach the front of the interface they disappear. This effect is reversed when the RSVP motion is in the backward direction, with new images appearing at the front and traveling to the back of the interface.

The images were introduced in a clockwise circular fashion. This means that the user could track the images around a circle and did not need to make any large jumps to view the next image. It was also thought that the progression of images from the back of the screen to the front would allow users to quickly discount an image and move on if it was not of interest.

Carousel

A carousel RSVP interface is shown in Figure 3.1. The display window was 450x300 pixels. Seven images were presented in the display window to the user. The images were presented in a carousel layout to the user, with the images getting larger as they reach the top. Due to the limitations of Tcl/Tk the images were not scaled in real time, but rather where only shown in the seven positions pictured. However as the RSVP motion was in progress the transition of the images either forward or backward was animated so the user would have feedback as to the current direction of the RSVP motion.

3.1.2 Control Sets

The three control sets used were selected, due to their different approaches and to cover Wittenburg et al.'s (2000b) criteria for an RSVP control set. These are that an image browser:

- Must support multiple speeds.
- Must be able to change direction.
- Must have an instant response time.

Figure 3.2: Carousel RSVP with the Keyboard Control Set

While all three control sets met the above criteria, they each employed a different approach. This was in order to provide insight into what is needed to provide a good control set for RSVP.

All of the control sets looped the images if the end of the image set was reached. This enabled users to quickly move back to the start of an image set if the current image displayed was near the end.

Keyboard Driven

An interface utilising the keyboard control set is shown in Figure 3.2. Navigation with the keyboard control set required no input with the mouse. The right arrow key either increased the RSVP presentation rate in the forward direction or decreased the RSVP presentation rate in the backward direction. The left arrow had the opposite effect. Both the forward and the backward direction had five separate speed settings. The down arrow instantly stoped the RSVP motion. However image selection still occurred with the mouse. This was due to the fact that in the floating and carousel RSVP modes, several images were displayed on screen at any one time. A possible disadvantage of this was that the change from keyboard to mouse control may result in slower image selection times for the keyboard control set.

An advantage of the keyboard control set is that it was possible to use while completely focusing on the RSVP display window. This may result in a lower miss rate, as users will not have to focus their attention on the control set.

Dial

The dial control set is shown in Figure 3.3. It was named dial because of its similarity with an analogue dial control. The control set consisted of several co-centric circles with an arrow running from the middle of the circles to the edge of the outermost circle. To initiate RSVP motion, users had to click anywhere inside the outside circle and then drag the mouse in either a clockwise or anti-clockwise fashion. This rotation caused the images to progress either in a forward or backward motion, depending on whether the circles made were clockwise or anti-clockwise respectively. While the dial was being manipulated, the arrow moved with the mouse to give in indication of which direction the images were moving. The RSVP presentation rate was dependent on the speed at which the circles were being made with the mouse.

The dial control is the only control set that put the RSVP presentation rate purely in the hands of the user. With the other control sets, the RSVP motion could continue without input from the user. While this gave the user more direct control, it may also have resulted in a higher physical demand.

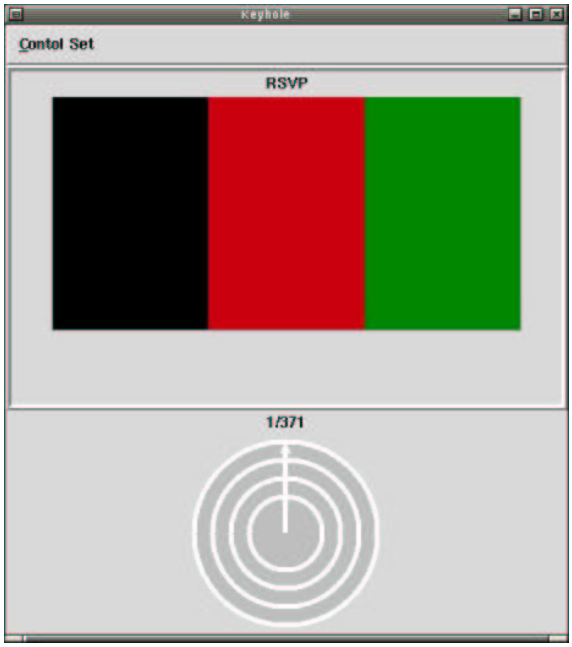


Figure 3.3: Keyhole RSVP with the Dial Control Set

Scroll Bar

An interface utilising the control bar set is shown in Figure 3.1. The user was presented with a scroll bar which had the slider centred at the middle. The slider could be dragged to the left and right of centre. Moving the slider right of centre initiated the RSVP motion in the forward direction, and moving the slider left of centre did so in the backward direction. If the slider was released, then it automatically returned to the centre and the RSVP motion was stopped. Using the middle mouse button users could go directly to a speed without having to grab the slider. Clicking anywhere on the scroll bar with the middle mouse button caused the slider to jump to that part of the scrollbar and the RSVP motion began. A click through effect could also be exploited by clicking on either side of the slider while it was centreed. This either incremented or decremented the image depending on which side of the slider was clicked.

By using a familiar device like the scroll bar, it was hoped that users would feel comfortable with this interface and find it intuitive to use.

Alternate Approaches

Several other approaches were also investigated. These included a mouse over arrow bar such as was implemented in Wittenburg et al's (2000b) online bookstore. However this interface was found to perform poorly and so was not included in the formal evaluation.

3.1.3 Speed Settings

Both the keyboard and scrollbar control sets allow a maximum RSVP presentation which the user cannot exceed. For both the carousel and keyhole modes, this value was initially based on the experimental results of Spence (2002). Their findings would equate to 25 images/sec for keyhole and 10 images/sec for carousel. However these values were for searching information and were found not to be suitable for browsing. The values were set at 15 images/sec for keyhole and floating and 13 images/sec for carousel. These speeds were determined in an informal evaluation with an expert user.

Figure 3.4: Simple Thumbnailed Interface

3.2 Thumbnail Browser

The implementation of a thumbnail browser is shown in Figure 3.4. The images were arranged on a grid and could be navigated through using a scroll bar.

Chapter 4

Evaluation

The evaluation compared our ten interfaces. Nine of these implemented RSVP and one was a traditional thumbnailed approach.

The comparison was primarily in terms of how long it subject took to retrieve an image using the interfaces. A subjective measurement of the interfaces was also undertaken through the NASA Task Load Index (TLX) worksheets.

4.1 Experimental Design

The experiment was a multi-factored within subjects anolysis of varience, as can be seen in Table 4.1. The factor 'Control Set' had three levels: Dial, Scroll Bar and Keyboard. The factor 'RSVP Mode' had three levels: Keyhole, Floating and Carousel. The best combination of RSVP mode and control set was then compared against the thumbnailed interface with a Paired T-Test.

		Control Set		
		Scroll Bar	Keyboard	Dial
	Keyhole	S1-12	S1-12	S1-12
RSVP Mode	Carousel	S1-12	S1-12	S1-12
	Floating	S1-12	S1-12	S1-12

Table 4.1: Experimental Design: Repeated Measures for RSVP Mode and Control Set.

4.1.1 Procedure

Subjects were asked to perform eleven image retrieval tasks, one for each of the RSVP interfaces and two for the thumbnailed interface. In total each RSVP mode and control set was tested three times. The order in which the subjects encountered the ten interfaces was varied to counter learning effect. In addition, six subjects encountered the RSVP interfaces in groups of control set, and six in groups of RSVP mode.

Each RSVP interface was only tested once, as it was the modes and the control sets that we were interested in rather than the specific interfaces. Having subjects use each interface multiple times would have extended the experiment to an hour on average. The experiment was quite demandeding for the subjects as it required them to be quite focused. It was felt that their performance would be impaired if the experiment took too long.

The same set of 371 images was used for each subject and for each retrieval task. Each task used a different target image to ensure that a subject did not get accustomed to looking for just one image. However, all the images were chosen to be 'recognisable'. For each task a different ordering of the images was chosen. This was so that subjects did not find images using their knowledge of the relative ordering of the images.

Figure 4.1: Before engaging in the image retrieval task, the target image was displayed to the subject

Due to the nature of the control sets, subjects could quickly navigate to the beginning and the end of the image set. Therefore the target image always lay within the middle third of the image set, to ensure the subject would have to spend some time using the interface. For each of the control sets and RSVP modes, the target image appeared once in a position from 124th to 165th, once in a position from 166th to 207th and once in a position from 208th to 249th. The positions were chosen so that the average position of the target image would average to the middle image for each control set and RSVP mode. This was the same for the thumbnailed browser tasks.

The target image was first presented to the user in a separate window, which is shown in Figure 4.1. Subjects were encouraged to spend at least ten seconds memorising the image and were allowed to take as much time as they liked. After the target image window was closed, the interface to be used appeared and the browsing task started. To complete the browsing task, subjects had to find the target image using the interface and select the image. A selection was made by clicking on the image with the mouse. In the case of the carousel and floating RSVP interfaces where several images were presented on screen, the subject could select any of the visible images. On selection of a correct image, a confirmation message was presented and the interface closed. Subjects were timed from when the interface appeared to the selection of the correct image. Timing was done automatically and was recorded to the nearest millisecond.

While carrying out the retrieval task, the target image was not available to the user. This is an image browser should be simple enough to allow the user to remember the target image (Combs & Bederson 1999). A browsing interface that causes a user to forget their task is a poor one.

Before using each of the interfaces, subjects were given a guided tutorial and were told how to use each interface. This included a full practice run where a browsing task was to be completed by the subject. Subjects were encouraged not to go on to the timed task until they were confident in using the interface.

To assess the stress and frustration felt by subjects, they were asked to fill in a NASA TLX worksheet (Hart & Staveland 1988) for each control set and RSVP mode and one for the thumbnailed interface. The NASA TLX allows ratings from 1 to 5 with 1 being better and has the following categories, Mental Demand, Physical Demand, Temporal Demand, Performance, Effort and Frustration Level.

After the evaluation subjects were encouraged to give extra comments and were asked whether they would ever consider using any of the RSVP interfaces in preference to the traditional thumbnailed approach.

Each evaluation lasted approximately 30 minutes.

4.1.2 Subjects

There were twelve subjects involved in the evaluation, all of whom were 4th year computer science students at the University of Canterbury. All twelve were expert computer users, of which half had approximately 30 minutes experience using RSVP based interfaces.

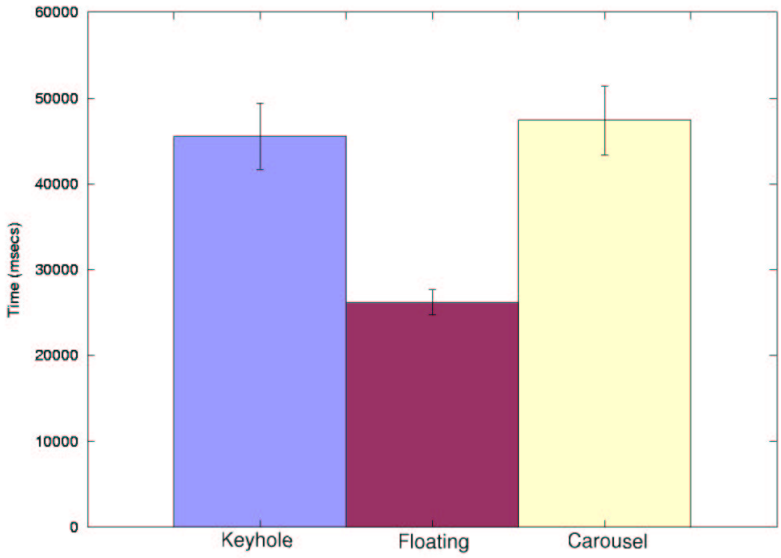


Figure 4.2: Mean Retrieval Times by RSVP Mode with Error Bars Showing Standard Error

4.2 Results

The evaluation was designed primarily to compare the three RSVP modes and the three control sets. Users also completed tasks with a thumbnailed image browser so as to provide a comparison.

4.2.1 RSVP Interface Comparison

Image Retrieval Time

The mean task completion time across all the RSVP interfaces was 39.7 (σ 22.0) seconds. Combs & Bederson (1999) found that users with the Thumbs+ image browser, a traditional thumbnailed approach, took a mean time of approximately 40 seconds to retrieve an image in an image set with 225 images. Comparatively our interfaces appear to be quite fast.

Looking first at the effect of RSVP mode, the mean task completion times for keyhole, floating and carousel were 45.5 (σ 23.3), 26.1 (σ 8.7) and 47.4 (σ 24.0) seconds. The mean task completion times by RSVP mode are shown in Figure 4.2. There was a significant difference between the modes ($F_{2,22}$=13.81, $p<0.001$). Application of the Tukey test produced a HSD (Honest Significant Difference) of 19.4 seconds. This indicates that floating RSVP was significantly faster than both carousel and keyhole. Carousel and keyhole were not significantly different to each other.

By control set the mean completion times for the dial, scrollbar and keyboard control sets were 37.5 (σ 14.9) , 43.4 (σ 31.0) and 38.1 (σ 16.6) seconds respectively. These results are summarised in Figure 4.3. There was no significant difference found between the control sets ($F_{2,22}$=1.24, p=0.309), all control sets performed equally well

Subjective Measures

The NASA TLX responses on the RSVP modes revealed several significant differences. These are shown in Figure 4.5(b). Mental demand was significantly different (Friedman Test, x_r^2=9.88, df=2, $p<.05$) with the mean responses for keyhole floating and carousel being 3.92 (σ 1.04), 2.42 (σ 0.64) and 3.33 (σ 0.85). Temporal demand was significantly different (Friedman Test, x_r^2=7.54, df=2, $p<.05$) with the mean responses for keyhole floating and carousel being 3.42 (σ 1.19), 2 (σ 0.58) and 2.83 (σ 0.99). Performance was significantly different (Friedman Test, x_r^2=6.16, df=2, $p<.05$) with the mean responses for keyhole floating and carousel being 3.08 (σ 0.86), 2.08 (σ 0.86) and 3 (σ 0.71). Frustration levels were significantly

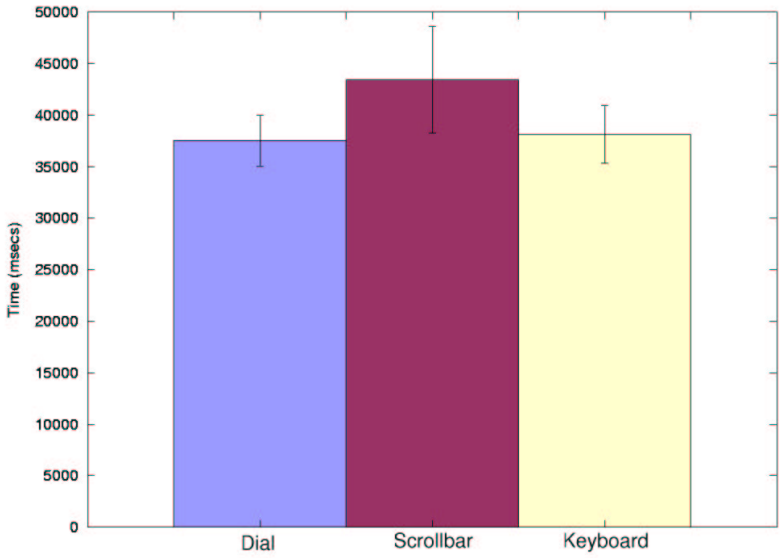


Figure 4.3: Mean Retrieval Times by Control Set with Error Bars Showing Standard Error

different (Friedman Test, x_r^2=6.79, df=2, p<.05) with the mean responses for keyhole floating and carousel being 3.17 (σ 1.21), 1.92 (σ 0.64) and 2.83 (σ 1.28).

Floating RSVP was found by users to be less demanding mentally and temporally than both keyhole and carousel. Subjects also felt they performed better with the floating interfaces and found them less frustrating. This is in line with our results for the mean retrieval times.

For the NASA TLX response on the control sets, there were two significant differences. A summary of these responses is shown in Figure 4.5(a). Physical demand was one (Friedman Test, x_r^2=7.54, df=2, p<.05), with the mean responses for dial, scrollbar and keyboard being 3.75 (σ 1.23), 2.67 (σ 0.94) and 2.25 (σ 1.01). Frustration was also significantly different (Friedman Test, x_r^2=4.63, df=2, p<.05) with the mean responses for dial scrollbar and keyboard being 3.33 (σ 1.18), 2.17 (σ 0.69) and 3.00 (σ 1.10).

Users found the dial control set the most physically demanding, and the scroll bar control set the least frustrating.

4.2.2 Thumbnailed Browser Comparison

As only one task was performed with each interface, a valid comparison could not be made bewtween a single RSVP interface and the thumbnailed interface. The comparison is therefore made between the mean results for the floating RSVP interfaces and the thumbnailed browser. This is as the floating interfaces were found to significantly faster than the keyhole and carousel interfaces.

The mean retrieval time for the thumbnailed interface was 42.3 (σ 30.2) seconds. While this was higher than the floating interfaces mean of 26.1 (σ 8.7) seconds, there was not a significant difference between the two means (p=0.0976). However, there were some significant differences in the NASA TLX responses.

The mean NASA TLX responses for both floating and thumbnailed can be seen in Figure 4.5(a). The results for physical demand are significantly different (Friedman Test, x_r^2=4.08, df=2, p<.05), with subjects finding the floating RSVP interfaces less physically demanding. Frustration levels were also significantly different (Friedman Test, x_r^2=4.08, df=2, p<.05) with subjects finding the floating interfaces less frustrating. Performance was also significantly different (Friedman Test, x_r^2=6.75, df=2, p<.05) with subjects feeling they performed better with the floating interfaces than the thumbnailed interfaces.

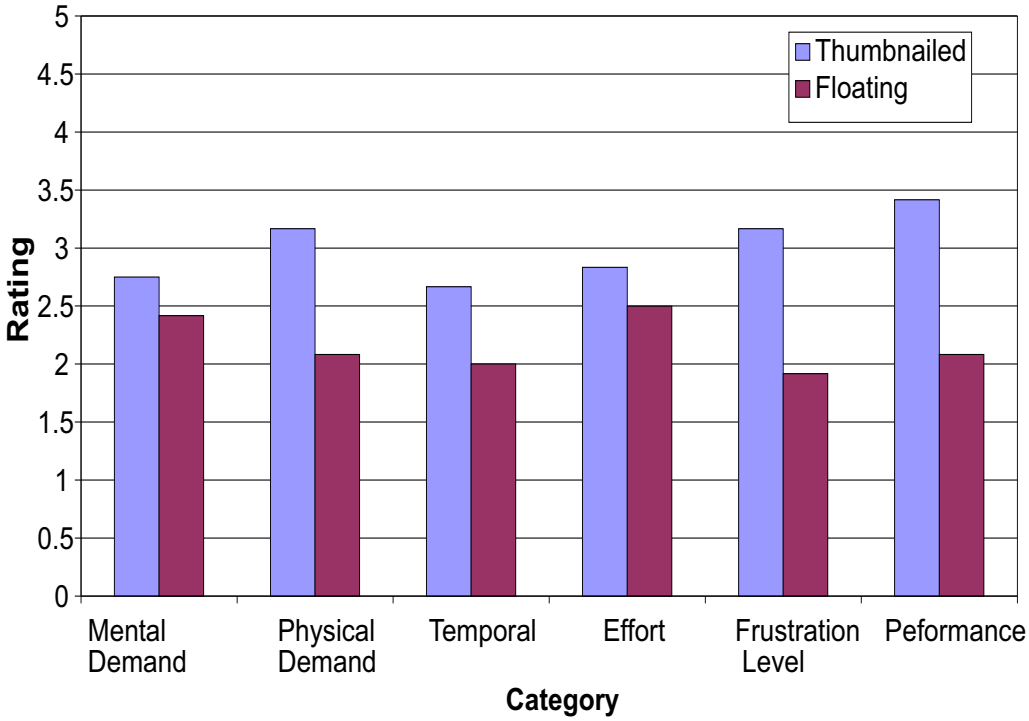


Figure 4.4: NASA TLX Comparison of Mean Responses for Floating and Thumbnailed

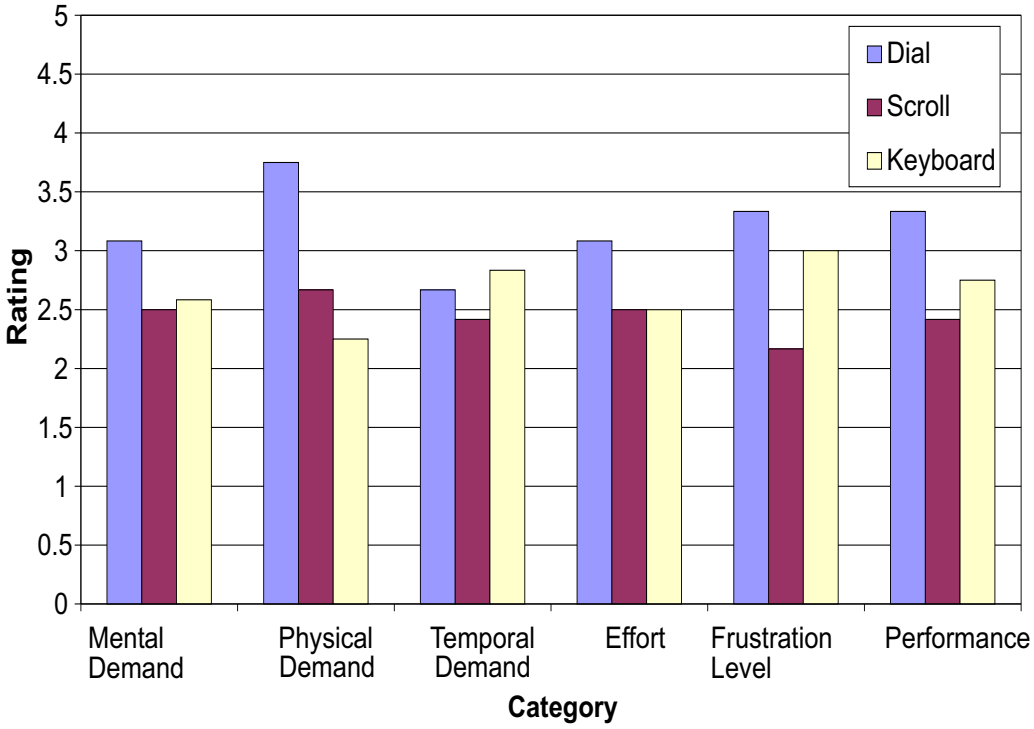
4.2.3 User Comments for All Interfaces

Many subjects commented that they found the keyhole interfaces quite stressful. They felt that if they lost concentration momentarily that they could miss the image they were looking for. With the floating and carousel interfaces, several images were on screen at any one time and subjects commented that this gave them more of a chance to evaluate and discard images. However one subject did comment that he enjoyed keyhole due to its simplicity. Several subjects did find that the carousel mode made them dizzy.

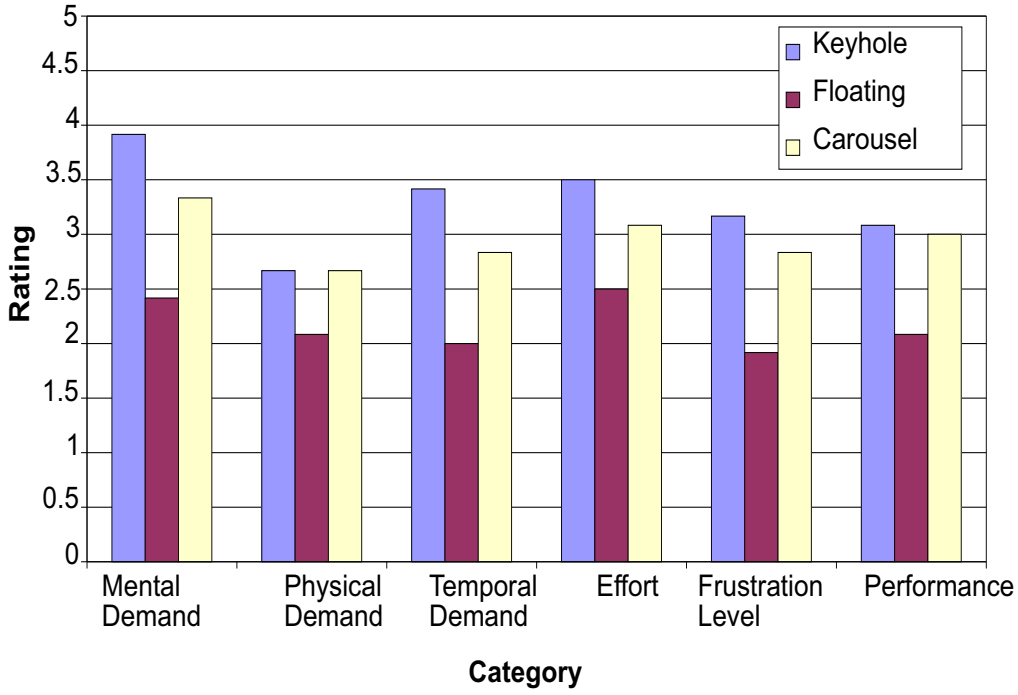
Several subjects found the dial control frustrating as they needed to concentrate both on the images and moving the dial in a regular fashion. However, one subject preferred the dial as he felt it gave him more control over the RSVP motion.

Several subjects also commented that they found it harder to remember the image they were looking for using the thumbnailed interface as opposed to the RSVP interfaces.

When asked if they would consider using a RSVP image browser as their regular image browser, ten of the twelve said they would.



(a) Ratings by Control Set



(b) Ratings by RSVP Mode

Figure 4.5: Mean NASA TLX Responses for RSVP Interfaces

Chapter 5

Discussion and Further Work

5.1 Experimental Concerns

For all of the selection tasks, a different image was used. While each of these images was chosen as 'recognisable', the choice was made by the author and the subjects may have found some of the images easier to recognise than others. It is hoped that this factor was balanced out across interfaces, so that none of the control sets or RSVP modes were biased due to their target images, but this is always a possibility. This potential bias could have been corrected by using the same image across all the interfaces.

In addition, the different maximum presentation rates of the RSVP modes could have been a confounding factor. However, this seems unlikely, as very few subjects actually used the maximum presentation rates, instead preferring to utilise a rate they felt comfortable with. This could become a major factor with expert users. A system for expert users would have to let them set their own maximum presentation rates.

One question of the evaluations validity concerns the use of flags for the image set. Users of an image browser would typically browse their own collection of digital media and would be more familiar with the collection. The choice of flags was motivated by the availability of high quality images many of which would be recognisable to the user. By using flags it was hoped that subjects would be familiar with the images and perform naturally.

5.2 Discussion of Results

In both the time taken to retrieve an image and subject preference, floating RSVP performed the best. Subjects generally felt more comfortable using higher presentation rates with the floating RSVP interfaces, which can explain the better retrieval times. The introduction of new images in a circular fashion could have resulted in a comfortable viewing pattern and could explain subject preference. In addition, although with floating RSVP the images changed size as they moved toward or away from the subject, they did not have much horizontal or vertical movement. This meant that the subject could easily track the image as it moved.

Several subjects commented that the carousel interfaces made them dizzy. One explanation for this is that the images moved a considerable distance on the screen, and subjects tracking an interesting image had to track it round the carousel which could have been quite tiresome. The tracking process is shown in Figure 5.2. A subject wishing to track an image would have to move the focus of their eyes from position 1 through to position 2, along the path of the carousel.

One flaw of the carousel interfaces is shown in Figure 5.2. New images entered the carousel at position 1 as shown in Figure 5.2. It was common for subjects to focus their attention here while looking for the target image. If the image proved interesting, they would occasionally track the image round the carousel until position 2, where the image left the display. From here subjects generally wished to focus their attention back at position 1 and the easiest way to do this was to 'jump' their eyes right back to position 1. Unfortunately this meant that they missed the six new images that had been displayed when their attention was focused on the interesting image. This problem is not encountered in the keyhole or floating interfaces.

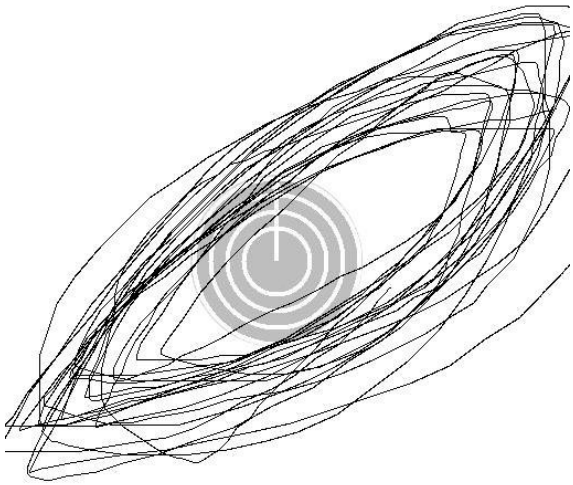


Figure 5.1: Typical mouse movements of a subject using the dial interface

As the keyhole interfaces only present one image on screen at a time to the subject, there is no chance of 'jumping' several images. With the floating RSVP, the next image is always close to the image being watched and so it is hard to 'jump' images.

Many subjects found the dial control set to be hard to control and hard to keep at a constant speed. The typical mouse movement of a subject with the dial control set is shown in Figure 5.1. We can see that subjects typically made ovals rather than circles with the mouse. This resulted in a quick movement when the mouse was closest to the center of the dial, and a slower movement when the mouse was furthest away from the center. This is understandable as subjects where more concerned with looking for the target image, than controlling their mouse movements.

There was no significant difference in the time taken to retrieve an image over the three control sets. This suggests that any control set that meets Wittenburg et al.'s (2000b) criteria for an RSVP control set should provide users the capability to effectively browse images using RSVP. However users preferences cannot be ignored and the scroll bar control set was found to be the least frustrating control set. The deciding factor into what control set to use would probably be the context in which the RSVP browser is used. For instance when applied to a typical mobile phone, the keyboard control set would be the only practical solution.

5.3 Further Work

Effect of 3D in RSVP Modes

Floating RSVP differed from both keyhole and carousel in that it displayed images using a 3D approach. This could have been a factor in its increased performance. Further work is required to investigate the effect of 3D metaphors in RSVP modes.

Cognitive Study of RSVP Modes

To date the science behind proposed RSVP modes has been limited. A study of RSVP modes from a cognitive perspective should return useful information. Looking at how the introduction of new images could be optimised to reduce eye strain would be useful, as would looking at the limits of human perception when deciding on maximum presentation rates.

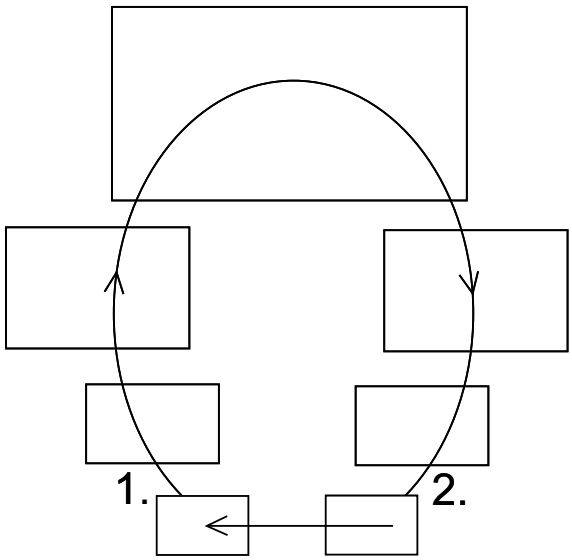


Figure 5.2: Occasionally subjects would ‘jump’ several images.

Expert User Evaluation

In developing the RSVP interfaces, the author became an expert user with them and found that he could find images effectively with all the interfaces, even at the highest presentation rates. An evaluation where users had more time to train and become familiar with the interfaces should reflect this experience with faster retrieval rates. An expert study may show the RSVP interfaces to be significantly faster than the traditional thumbnailed approach.

This is backed up by the results found by Ahlberg & Shneiderman (1994) and their alphaslider. Expert users were able to select items from a list much faster than novice users.

Application to Small Screen Devices

A typical small screen device such as the Compaq Pocket PC has a resolution of 320x240 (Wobbrock et al. 2002). It is believed that our RSVP interfaces would easily scale to this size with no performance decrease from their current average size of 460x400. The RSVP interfaces are much less dependent on screen real estate than a thumbnailed approach, and so should perform comparatively much better on small screen devices.

The control sets could be adapted to a small screen device. While the options on todays cellphones would probably limit the choice of control sets to keyboard, the input options on other small screen devices such as PDA’s would allow us to implement all the control sets. Implementing the keyboard control would be trivial, as it only requires three keys. The scroll bar approach could be harder, but this could either be approximated by a keyboard driven approach, or with stylus input. The dial approach would be hardest to implement in software, as the graphical representation of the dial takes up a fair amount of screen real estate. However, a physical dial that the user could turn in a similar fashion to a volume control on a stereo could prove to be a success. A physical dial would also get past some of the frustrations that users had with our dial implementation as they would no longer have to concentrate on making accurate mouse movements.

Small screen devices with the capacity to store, send and receive pictures are set to both increase in popularity and power. An evaluation into the effectiveness of using an RSVP image browser in the context of a small screen device would be desirable.

Chapter 6

Conclusion

In this report, a new image browsing technique using RSVP was presented and evaluated. Three control sets were developed, dial, scroll bar and keyboard, and tested with three RSVP modes, keyhole, floating and carousel. An evaluation was conducted to compare the effectiveness of the modes and the control sets. Both the time taken to retrieve an image and several measures of user satisfaction were recorded.

The evaluation showed that RSVP mode, rather than control set, had a greater bearing on effectiveness. Subjects were significantly faster when using interfaces implementing floating RSVP than when they used keyhole and carousel based interfaces. In addition, floating was also the most preferred RSVP mode. There was no significant difference in the image retrieval times with the three control sets, although subjects favoured the scroll bar control set. When the floating RSVP interfaces where compared to a traditional thumbnailed browser they performed faster, although not significantly so.

RSVP had already been shown to have potential in the general field of information navigation, and our evaluation showed that it has potential in the field of image browsers. In situations where screen real estate is limited, an RSVP image browser could be particularly useful, due to RSVP's space-time tradeoff.

Acknowledgments

I'd just like to thank some people for their help during the year. Thanks to Andy Cockburn for his endless enthusiasm and to Stacy Mickelbart for helping keep my ramblings legible. Cheers to the honours crew for keeping this year fun, gots to love those indoor cricket sessions. Big ups to the Suva St crew for a generally wicked year. Lastly I'd like to thank Cat for her love and support.

Bibliography

Ahlberg, C. & Shneiderman, B. (1994), The alphaslider: A compact and rapid selector, *in* 'Proceedings of CHI-94', Boston, MA, pp. 365–371.

Bederson, B. (2001), Photomesa: A zoombable image browser using quantum treemaps and bubblemaps, *in* 'The 14th Annual ACM Symposium on User Interface Software and Technology (UIST'01)', ACM Press, pp. 71–80.

Combs, T. T. & Bederson, B. B. (1999), Does zooming improve image browsing, *in* 'Proceedings of the Fourth ACM Conference on Digital libraries', ACM Press, pp. 130–137.

de Bruijn, O. & Spence, R. (1999), Rapid serial visual presentation: a space-time trade-off in information presentation, *in* 'Proceedings of the Working Conference on Advanced Visual Interfaces', ACM Press, pp. 189–192.

de Bruijn, O. & Spence, R. (2002), Patterns of eye gaze during rapid serial visual presentation, *in* 'Proceedings of AVI2002', ACM Press.

de Bruijn, O., Spence, R. & Yin, M. (2001), 'Rsvp browser: Web browsing on small screen devices'. *citeseer.nj.nec.com/debruijn01rsvp.html

Derthick, M. (2002), Interfaces for palmtop image search, *in* 'Proceeding of the Second ACM/IEEE-CS Joint Conference on Digital Libraries', ACM Press, pp. 340–341.

Forster, K. L. (1970), Visual perception of rapidly presented word sequences of varying complexity, *in* 'Perception & Psychophysics', pp. 215–221.

Graham, A., Garcia-Molina, H., Paepcke, A. & Winograd, T. (2002), Time as essence for photo browsing through personal digital libraries, *in* 'Proceedings of the Second ACM/IEEE-CS Joint Conference on Digital libraries', ACM Press, pp. 326–335.

Hart, S. & Staveland, L. (1988), Development of NASA-TLX (task load index): Results of empricial and theoretical research, *in* P. Hancock & N. Meshkati, eds, 'Human Mental Workload', Elsevier Science, pp. 139–183.

Krishnamachari, S. (1999), Image browsing using hierarchical clustering, *in* 'Fourth IEEE Symposium on Computers and Commications ISCC'99', IEEE, pp. 301–307.

Mikkel, A. (1992), *Digital Photography*, Random House, p. 9.

Rahman, T. & Muter, P. (1999), Designing an interface to optimize reading with small display windows, *in* 'Human Factors', Human Factors and Ergonomics Society.

Rodden, K., Basalaj, W., Sinclair, D. & Wood, K. (2001), Does organisation by similarity assist image browsing?, *in* 'Special Interest Group on Human-Computer Interaction SIGCHI'01', ACM Press, pp. 190–197.

Spence, R. (2002), 'Rapid, serial and visual: a presentation technique with potential'.

Tatemura, J. (2000), Dynamic label sampling on fisheye maps for information navigation, *in* 'Proceedings of the Working Conference on Advanced Visual Interfaces (AVI'2000)', ACM Press, pp. 238–241.

TriVista Technologies (2002), 'TriVista Technologies', *http://www.trivista.com/*.

Tse, T., Marchionini, G., Ding, G., Slaughter, L. & Komlodi, A. (1998), Dynamic key frame presentation techniques for augmented video browsing, *in* 'Proceedings of the Working Conference on Advanced Visual Interfaces (AVI'1998)', ACM Press, pp. 185–194.

Wittenburg, K., Ali-Ahmed, W., LaLiberte, D. & Lanning, T. (1998), Rapid-fire image previews for information navigation, *in* 'Proceedings of the Working Conference on Advanced Visual Interfaces', ACM Press, pp. 76–82.

Wittenburg, K., Chiyoda, C., Heinrichs, M. & Lanning, T. (2000b), Browsing through rapid-fire imaging: Requirements and industry initiatives, *in* 'Proceedings of Electronic Imaging'2000'.

Wittenburg, K., Nicol, J., Paschetto, J. & Martin, C. (2000a), Browsing with dynamic key frame collages in web-based entertainment video services, *in* 'International Conference On Multimedia Computing and Systems', IEEE.

Wobbrock, J. O., Forlizzi, J., Hudson, S. E. & Myers, B. A. (2002), Webthumb: Interaction techniques for small-screen browsers, *in* 'Proceedings of the 15th Annual ACM Symposium on User Interface Software and Technology', ACM Press, pp. 205–208.